

## **CHAPTER 3**

### **RESEARCH METHODOLOGY**

#### **3.1 Research Instrument**

##### **1. Software**

- a) The operating system that will used is Ubuntu 14.04 LTS
- b) Netbeans 8.2 that obtained through <https://netbeans.org/>

##### **2. Hardware**

- a) The processor that will used is AMD A4-3330MX APU with Radeon(tm) HD Graphics
- b) 500 GB hard driver
- c) 2GB RAM

#### **3.2 Data Collection**

The training data and the testing data are obtained through web scraping process using a tool which is Web Scraper Google Chrome Extensions (source : <https://www.webscraper.io/>). The data will obtained in form .csv format then imported into database MYSQL. The data are taken from <https://www.bukukita.com/> consist 60 data for the training data.

#### **3.3 Research Procedure**

##### **1. Literature Study**

Collecting 10 journals related to data mining, LVQ algorithm an TF-IDF then make analysis how the problems could solved using those implementation.

## 2. Data Collection

Collecting training data and testing data of the books using Web Scraper Google Chrome Extensions (source : <https://www.webscraper.io/>) to web scrapping the data.

## 3. System Design

The system design should contain how LVQ method work for classify book category. The step design will be described with a flow chart. This process will help to keep how the procedure work is.

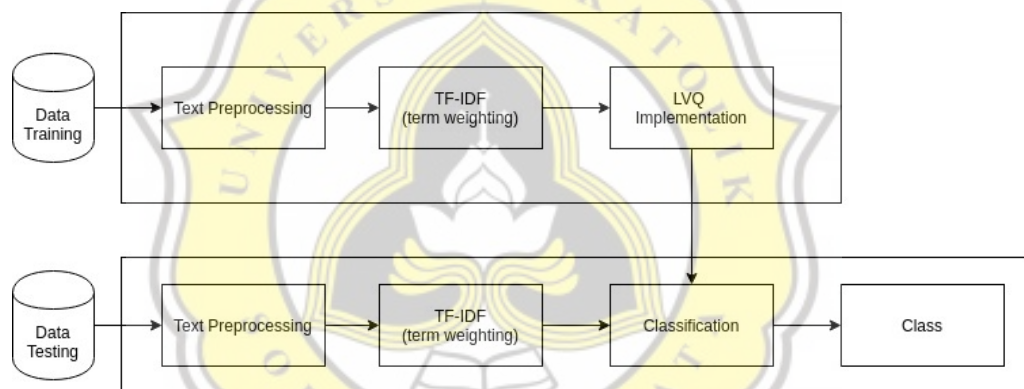


Illustration 3.1: Program System Design

## 4. Coding & Testing

After data already obtained from web scraping process using a tool which is Web Scraper Google Chrome extensions. The first thing to do is import the data into MYSQL database. Next step is the preprocessing stage which is contain casefolding, tokenization, stopwords and stemming. After that it will features extraction using TF-IDF and the result of the weighting will be converted into vector form. Then the program will implement LVQ algorithm for the training data using differences of epoch and learning rate. This project program codes is using Java programming language. In testing process, testing data will also proceed same procedure with training data but after TF-IDF stage, testing data will be classified using the LVQ implementation from training data before. In the testing process can be found how the accuracy of LVQ algorithms.

## 5. Report

Write the report about the classification using Learning Vector Quantization algorithm in this project. The report contains from initial process to the result and conclusion about how Learning Vector Quantization algorithm implementation to classify book categories.

